4. Topological phases-d

The simplest way of realizing the QSH state is by superposing two QH systems with opposite spin, or put in another way, a QH system with its time-reversed counterpart. Use for example the Haldane model to construct the Hamiltonian

$$\mathcal{T}H\mathcal{T}^{-1} = H' \neq H \;\rightarrow\; H_{doubled} = \begin{pmatrix} H_\uparrow & 0 \\ 0 & H_\downarrow \end{pmatrix} = \begin{pmatrix} H_{Haldane}(\phi = -\pi/2) & 0 \\ 0 & H'_{Haldane}(\phi = \pi/2) \end{pmatrix}$$

The superposition of the down-spin QH subsystem and the up-spin QH subsystem results in $v = v_\uparrow + v_\downarrow = 0$. While the sum vanishes due to time-reversal symmetry, the difference $v_s = v_\uparrow - v_\downarrow$ is non-zero and can be used to define a "spin Chern number", v_s, that counts number of time-reversed (Kramer's) pairs of edge modes. In this simple example s_z is conserved, which may not be the case in a real system where spin-coupling terms will mix the two sub-systems. The question is what happens when s_z is no longer a good quantum number.

Spin-orbit coupling

In previous discussions, we ignored the spin degrees of freedom. In reality, we have spin up and spin down electrons. If we consider free electrons moving in 2D (without lattice), the Hamiltonian is:

$$H = \sum\nolimits_k \frac{k^2}{2m} c_{k\uparrow}^\dagger c_{k\uparrow} + \sum\nolimits_k \frac{k^2}{2m} c_{k\downarrow}^\dagger c_{k\downarrow}$$

Here the kinetic energy for a particle with momentum k and spin σ is $\epsilon_\sigma(k) = k^2/2m$. The kinetic energy doesn't depend on the spin degrees of freedom, so we say that: spin degrees of freedom and orbit degrees of freedom (i.e. the momentum k) decouple.

This Hamiltonian (without spin-orbit coupling) is an ideal case. In reality, spin and orbit motions are couple together.

More general cases:

Because we have two spin species, the Hamiltonian of the kinetic energy is in general a 2×2 matrix:

$$H = \sum\nolimits_k \begin{pmatrix} c_{k\uparrow}^\dagger & c_{k\downarrow}^\dagger \end{pmatrix} \mathcal{H} \begin{pmatrix} c_{k\uparrow}^\dagger \\ c_{k\downarrow}^\dagger \end{pmatrix}$$

Because the Hamiltonian is a Hermitian operator, $\mathcal{H}$ must be a Hermitian 2×2 matrix. For any Hermitian 2×2 matrix, we can write it in terms of identity and Pauli matrices.

$$\mathcal{H} = \mathcal{H}_0 I + \mathcal{H}_x \sigma_x + \mathcal{H}_y \sigma_y + \mathcal{H}_z \sigma_z$$

Here, the Pauli matrices are the true spin Pauli matrices, and they do tell us the spin of an electron.

For the case we consider above, $\mathcal{H}_x = \mathcal{H}_y = \mathcal{H}_z = 0$, we only have $\mathcal{H}_0 = k^2/2m$. There is no Pauli matrix in the Hamiltonian, which means that the spin degrees of freedom play no role in the kinetic energy. In general, the $\mathcal{H}_x, \mathcal{H}_y, \mathcal{H}_z$ can be nonzero, and thus the spin and orbital motion couples together.

One example:

$$\mathcal{H}(k) = \begin{pmatrix} \frac{k^2}{2m} + \alpha k & 0 \\ 0 & \frac{k^2}{2m} - \alpha k \end{pmatrix}$$

More realistic model: the Rashba spin-orbital coupling:

$$\mathcal{H} = \frac{k^2}{2m} I + \alpha(\sigma \times k) \cdot \hat{z} = \frac{k^2}{2m} I + \alpha\sigma_x k_y - \alpha\sigma_y k_x = \begin{pmatrix} \frac{k^2}{2m} & \alpha k_y + i\alpha k_x \\ \alpha k_y - i\alpha k_x & \frac{k^2}{2m} \end{pmatrix}$$

This Rashba term, $(\sigma \times k) \cdot \hat{z}$, breaks the symmetry of $z \to -z$, and therefore, it will only arise in systems without $z \to -z$ symmetry, e.g. the interface between two different materials or the surface of a material. For the interface, because the two materials have different electron density, electron will redistribute near the interface (moving from on material to the other). This induces an E field perpendicular to the interface along the z direction. This E field breaks the $z \to -z$ symmetry and thus can generate a nonzero Rashba term. Here, the coupling strength α is determined by the strength of this E field.

Another example: the Dresselhaus spin-orbital coupling:

$$\mathcal{H}_{\text{Dresselhaus}} = \beta\left[\sigma_x k_x\left(k_x^2 - k_y^2\right) + \sigma_y k_y(k_z^2 - k_x^2) + \sigma_z k_z\left(k_x^2 - k_y^2\right)\right]$$

Time-reversal symmetry

Typically, spin-orbit couplings are described by terms like $\sum_{ij} c_{ij}\sigma_i k_j^{2n+1}$. The spin-orbit coupling PRESERVES the time-reversal symmetry.

$$\mathcal{T}\sigma \to -\sigma \quad and \quad \mathcal{T}k \to -k$$

$$\mathcal{T}\sigma_i k_j^{2n+1} = (-\sigma_i)\left(-k_j^{2n+1}\right) = (-1)^{2n+2}\sigma_i k_j^{2n+1} = \sigma_i k_j^{2n+1}$$

Spin-orbital coupling and magnetic field

Starting from the model we considered above

$$\mathcal{H}(k) = \begin{pmatrix} \frac{k^2}{2m} + \alpha k & 0 \\ 0 & \frac{k^2}{2m} - \alpha k \end{pmatrix}$$
$$= \begin{pmatrix} \frac{1}{2m}(k^2 + 2\alpha m k + m^2\alpha^2) - m\alpha^2/2 & 0 \\ 0 & \frac{1}{2m}(k^2 - 2\alpha m k + m^2\alpha^2) - m\alpha^2/2 \end{pmatrix}$$
$$= \begin{pmatrix} \frac{(k+\alpha m)^2}{2m} & 0 \\ 0 & \frac{(k-\alpha m)^2}{2m} \end{pmatrix} - m\alpha^2/2I$$

What does this Hamiltonian looks like? We review the situation of electrons moving in some magnetic field. But spin up and down electrons feels opposite magnetic field.

$$H = \frac{(k - eA)^2}{2m}$$

We can see that for spin up electrons, they are coupled to an "magnetic field" with $A_\uparrow = -\alpha m/e$. For spin down electrons, it is coupled to a magnetic field with $A_\downarrow = \alpha m/e$. $A_\uparrow = -A_\downarrow$, so $B_\uparrow = -B_\downarrow$. Notice that this is just a toy model for demonstration purposes. Strictly speaking, one cannot define a vector potential in 1D because $B = \nabla \times A$, but one cannot define cross product in 1D. So one need a 2D or 3D system to really see the analogy between B field and spin-orbit couplings.

Topological insulator from spin-orbit effect

- We can get quantum Hall insulators using B field (uniform or staggered)
- Spin-orbit effect is really similar to external B field, $\vec{B}(r)$ for spin up electron and $-\vec{B}(r)$ for spin down electron.

Can we use spin-orbit effect to get topological insulators? We can do that, but there are two things different from IQH.

Example: Haldane's model ×2=Kane-Mele Model

- We consider two copies of Haldane's model.
- One for spin up and one for spin down
- The complex hoppings for spin up and down electrons have opposite phase.
    - Spin-up: quantum Hall state with C=1
    - Spin-down: quantum Hall state with C=-1
    - The whole system still has C=0 (no Hall effect), but this is an insulator with metallic edge states.

Haldane model:

$$H = H_{NN} + H_{NNN} = \sum\nolimits_k \begin{pmatrix} a_k^\dagger & b_k^\dagger \end{pmatrix} \begin{pmatrix} \mathcal{H}_{11}(k) & \mathcal{H}_{12}(k) \\ \mathcal{H}_{21}(k) & \mathcal{H}_{22}(k) \end{pmatrix} \begin{pmatrix} a_k \\ b_k \end{pmatrix}$$

$$\mathcal{H}_{12}(k) = -t\left(e^{-i\overrightarrow{k'}\cdot\overrightarrow{e_1}} + e^{-i\overrightarrow{k'}\cdot\overrightarrow{e_2}} + e^{-i\overrightarrow{k'}\cdot\overrightarrow{e_3}}\right)$$

$$\mathcal{H}_{21}(k) = \mathcal{H}_{12}(k)^* = -t\left(e^{i\overrightarrow{k'}\cdot\overrightarrow{e_1}} + e^{i\overrightarrow{k'}\cdot\overrightarrow{e_2}} + e^{i\overrightarrow{k'}\cdot\overrightarrow{e_3}}\right)$$

$$\mathcal{H}_{11}(k) = -2t'[\cos(k\cdot v_1 - \phi) + \cos(k\cdot v_2 - \phi) + \cos(k\cdot v_3 - \phi)]$$

$$\mathcal{H}_{22}(k) = -2t'[\cos(k\cdot v_1 + \phi) + \cos(k\cdot v_2 + \phi) + \cos(k\cdot v_3 + \phi)]$$

Kane-Mele model:

$$H = \sum\nolimits_k \begin{pmatrix} a_{k\uparrow}^\dagger & b_{k\uparrow}^\dagger & a_{k\downarrow}^\dagger & b_{k\downarrow}^\dagger \end{pmatrix} \begin{pmatrix} \mathcal{H}_{11}(k) & \mathcal{H}_{12}(k) & 0 & 0 \\ \mathcal{H}_{21}(k) & \mathcal{H}_{22}(k) & 0 & 0 \\ 0 & 0 & \mathcal{H}_{11}(k) & \mathcal{H}_{21}(k) \\ 0 & 0 & \mathcal{H}_{12}(k) & \mathcal{H}_{22}(k) \end{pmatrix} \begin{pmatrix} a_{k\uparrow}^\dagger \\ b_{k\uparrow}^\dagger \\ a_{k\downarrow}^\dagger \\ b_{k\downarrow}^\dagger \end{pmatrix}$$

This is an ideal model, assuming there is no hybridization between spin up and down electrons (the low-left and upper-right corner of $\mathcal{H}$ are assumed to be zero). But in any real system, this is not the case. One could get extra terms which hybridize the spin up and down electrons (extra terms in the Hamiltonian). In the presence of those extra terms, Kane and Mele found that the parity of the edge states (odd or even) is topologically invariant, as long as the time-reversal symmetry is preserved.

The Kane-Mele model

Continuum Version

We chose the same gauge and lattice orientation as in the original Kane and Mele papers. As such, we need to rotate the graphene lattice by $\pi/2$ compared with the one we used for the Haldane model: $k_x \to k_y$ and $k_y \to -k_x$. The $\mathbf{K}, \mathbf{K}'$ Haldane Hamiltonians become

$$H_{\mathrm{K+k}} = -3t'\cos(\varphi) - \frac{3}{2}t\left(k_x\sigma_x + k_y\sigma_y\right) + \left(m - 3\sqrt{3}t'\sin(\varphi)\right)\sigma_z$$

$$H_{\mathrm{K'+k}} = -3t'\cos(\varphi) - \frac{3}{2}t\left(-k_x\sigma_x + k_y\sigma_y\right) + \left(m + 3\sqrt{3}t'\sin(\varphi)\right)\sigma_z$$

We begin with a low-energy description of graphene by expanding the wavefunction around the points $\mathbf{K}, \mathbf{K}'$.We can mix the wavefunction around $\mathbf{K}, \mathbf{K}'$ in a four-component spinor:

$$\Psi(r) = \Big(\big(u_A(\mathbf{K}), u_B(\mathbf{K})\big), \big(u_A(\mathbf{K}'), u_B(\mathbf{K}')\big)\Big)\psi(r)$$

Assume no Haldane phase ($\varphi = 0$), and denote the Dirac velocity $v_F = -\frac{3at}{2}$ with $t < 0$. The low-energy description of graphene around each Dirac node is then given by

$$H_0 = -i\hbar v_F \psi^\dagger \big(\sigma_x \tau_z \partial_x + \sigma_y \partial_y\big)\psi$$

where σ acts on the sublattice index and $\tau_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ acts in the $\mathbf{K}, \mathbf{K}'$-points space. In matrix notation,

$$H_0 = -i\hbar v_F \psi^\dagger \begin{bmatrix} \sigma_x \partial_x + \sigma_y \partial_y & 0 \\ 0 & -\sigma_x \partial_x + \sigma_y \partial_y \end{bmatrix} \psi$$

where the two Hamiltonian blocks correspond to low energies around the points $\mathbf{K}, \mathbf{K}'$. In this low energy, we now want to look at the possible terms that can open a gap. As we know, σ_z at both $\mathbf{K}$ and $\mathbf{K}'$ can open a gap that is odd under inversion. The other term, the Haldane mass, is $\sigma_z \tau_z$; it provides for a mass that changes sign between the points $\mathbf{K}$ and $\mathbf{K}'$. The term is even under inversion (which changes both sublattices and $\mathbf{K}$ with $\mathbf{K}'$), but it is odd under $\mathcal{T}$.

Adding Spin

The Hamiltonian above is for spinless fermions. We now want to add spin to the system, which doubles it. For each spin $\uparrow, \downarrow$, the Hamiltonian becomes an 8×8 matrix if we include all spin, points $\mathbf{K}$ and $\mathbf{K}'$, and sublattice *A*, *B* structure in a single low-energy matrix. This is the low-energy gapless graphene Hamiltonian with spin.

Several gaps are possible: the boring ones are just doubled versions of the gaps that we encountered in the spinless version. If the gaps do not make use of the new spin structure, we are bound to obtain just doubled versions of the Haldane model. We have seen that $m\sigma_z + \tau_z \sigma_z 3\sqrt{3}t \sin(\varphi)$ does the job. However, this is not time-reversal symmetric for $\varphi \neq 0, \pi$ and m alone opens trivial gaps with $C = 0$.

With added spin, another gap-opening term is possible. Kane and Mele found that the only symmetry-allowed SOC terms in graphene open a small bulk energy gap which is topologically nontrivial. The SOC in graphene preserves spin-rotation symmetry around the *z* axis. Thus, s_z is a good quantum number and the system decomposes into two independent spin channels. Remarkably, each spin channel realizes a Chern insulator on its own with nontrivial but opposite Chern numbers $C_{\uparrow,\downarrow} = \pm 1$. This term could be

$$H_{\mathrm{SO}} = \lambda_{\mathrm{SO}} \psi^\dagger \sigma_z \tau_z s_z \psi$$

where s_z is the electron spin. The s_i matrices act on spin space, the σ_i act on the *A*, *B* sublattice space, and the τ_i act on the low-energy $\mathbf{K}, \mathbf{K}'$ space (valley degree of freedom). This term respects TR symmetry, as we show shortly. Kane and Mele also considered a spin symmetry breaking Rashba term, which can arise due to a perpendicular electric field or interaction with a substrate. For this broken mirror symmetry, we can have a term of the Rashba type, $(\mathbf{s} \times \mathbf{p}) \cdot \hat{z}$:

$$H_{Rashba} = \lambda_{Rashba} \psi^\dagger \left(\sigma_x \tau_z s_y - \sigma_y s_x \right) \psi$$

Let us now analyze the TR properties of these terms. Because we have added spin to the system, the TR operator has changed to $T = -is_y K$. The Hamiltonian at $\mathbf{K}$ without spin-orbit coupling can be written as

$$H_0(\mathbf{K} + \mathbf{k}) = \psi^\dagger \begin{bmatrix} h_\uparrow(\mathbf{k}) & 0 \\ 0 & h_\downarrow(\mathbf{k}) \end{bmatrix} \psi$$

where $h_\uparrow = h_\downarrow = k_x \sigma_x + k_y \sigma_y$ are the spinless graphene Hamiltonians around $\mathbf{K}$. We now want to add a constant in terms of $\mathbf{K} + \mathbf{k}$ that can open a gap. To open a gap, any extra added Hamiltonian ΔH must anticommute with H_0. If the added term commutes with any of the matrices in the Hamiltonian, then it will not open a gap: for example, $k_x \sigma_x$ and $a\sigma_x$ commute — an $a\sigma_x$-term does not open a gap but rather shifts the gapless point $k_x \to k_x - a$. Hence, the terms allowed that can open a gap must anticommute with σ_i because $H_0(\mathbf{K} + \mathbf{k}) = I_s \times k_i \sigma_i$ where I_s is the identity matrix in spin space. Hence, the only term allowed must contain σ_z in *A*, *B* sublattice space. Next, we need a spin dependent (s) part to couple the two copies of the Haldane model. Let us try for the $\mathbf{K}$ point:

$$\Delta H(\mathbf{K} + \mathbf{k}) = s_z \sigma_z = \begin{bmatrix} \sigma_z & 0 \\ 0 & -\sigma_z \end{bmatrix}$$

which gives us different gaps [with different "sign(*m*)"] for the two spins. How do we now add the valley degree of freedom (τ) in order to make it time-reversal invariant? In matrix form, the $\mathcal{T}$ operator is

$$\mathcal{T} = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix} K$$

with *K* complex conjugation and *I* the identity 2×2 matrix. For the Bloch Hamiltonian to be $\mathcal{T}$ invariant, we need

$$H(-\mathbf{k}) = \mathcal{T} H(\mathbf{k}) \mathcal{T}^{-1} = \begin{bmatrix} h_\downarrow(\mathbf{k})^* & 0 \\ 0 & h_\uparrow(\mathbf{k})^* \end{bmatrix}$$

The initial gapless Hamiltonian, $H_0(\mathbf{k})$, defined over the whole BZ is obviously TR invariant~~,~~. Time reversal relates the low-energy Hamiltonian at $\mathbf{K}$ with that at $\mathbf{K}'$ by flipping the spin:

$$\Delta H(-\mathbf{K} - \mathbf{k}) = \Delta H(\mathbf{K}' - \mathbf{k}) = -s_z \sigma_z = \begin{bmatrix} -\sigma_z & 0 \\ 0 & \sigma_z \end{bmatrix}$$

That means the term $\propto s_z\sigma_z$ transforms as

$$\mathcal{T}s_z\sigma_z\mathcal{T}^{-1} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}\begin{pmatrix} -\sigma_z & 0 \\ 0 & \sigma_z \end{pmatrix}\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix} = \begin{pmatrix} -\sigma_z & 0 \\ 0 & \sigma_z \end{pmatrix} = -s_z\sigma_z$$

Under time reversal, $\mathbf{K} \to \mathbf{K}'$. Hence, we need the gap opening term in $\mathbf{K}'$ to be $\mathcal{T}s_z\sigma_z\mathcal{T}^{-1} = -s_z\sigma_z$ to have $\mathcal{T}H(\mathbf{k})\mathcal{T}^{-1} = H(-\mathbf{k})$.

Thus, we have obtained the form of the low-energy gap-opening Hamiltonian around the points $\mathbf{K}$ and $\mathbf{K}'$ that preserves $\mathcal{T}$. Around $\mathbf{K}$, it is $s_z\sigma_z$. Around $\mathbf{K}'$, it is $-s_z\sigma_z$. If we consider the eight-component spinor that contains both spin, lattice, and $\mathbf{K}$, $\mathbf{K}'$ quantum numbers, then we conclude that **the full gap opening term should be of the form (a TR invariant mass)** is indeed

$$H_{KM} = \lambda_{\mathrm{SO}} s_z\sigma_z\tau_z$$

which a fully gap can be open. Moreover, H_{KM} is time-reversal invariant. This term can obviously arise only from spin-orbit coupling because it couples the spin with the momentum quantum numbers. Such a term opens a gap in the spectrum:

$$E_k = \pm\sqrt{\left(\hbar v_f k\right)^2 + \lambda_{\mathrm{SO}}^2}$$

Each band is doubly degenerate. We can write the tight-binding model as

$$H = t\sum_{\langle ij\rangle} c_i^\dagger c_j + i\,\lambda_{\mathrm{SO}} \sum_{\langle\langle ij\rangle\rangle} v_{ij} c_i^\dagger s_z c_j$$

which is just two copies of the Haldane model.

We now delve further into the physical properties of the doubled Haldane model by first separately analyzing the physics of the model for spin ↑ and ↓.

For spin ↑, the Hamiltonians at $\mathbf{K}$ and $\mathbf{K}'$ are

$$H_{\mathrm{K+k}} = k_x\sigma_x + k_y\sigma_y + \lambda_{\mathrm{SO}}\sigma_z$$

$$H_{\mathrm{K'+k}} = -k_x\sigma_x + k_y\sigma_y - \lambda_{\mathrm{SO}}\sigma_z$$

This is the $m = 0$ limit of the Haldane model studied in the last section. The Haldane mass term is λ_{SO} in this case. As such, in this regime, from our analysis of the Haldane model in the previous section, we know that the spin ↑ has a Hall conductance equal to 1.

For spin ↓, the Hamiltonians at $\mathbf{K}$ and $\mathbf{K}'$ are

$$H_{\mathrm{K+k}} = k_x\sigma_x + k_y\sigma_y - \lambda_{\mathrm{SO}}\sigma_z$$

$$H_{\mathrm{K'+k}} = -k_x\sigma_x + k_y\sigma_y + \lambda_{\mathrm{SO}}\sigma_z$$

This is the $m = 0$ limit of the Haldane model studied previously, but with an opposite Haldane term from the spin ↑ Hamiltonian. As such, from our analysis of the Haldane model, we know that the spin ↓ has a Hall conductance equal to −1.

The Hamiltonian we derived is just two decoupled Haldane Hamiltonians, one for spin up and one for spin down with opposite Hall conductances. To couple the two Hamiltonians, we can also add a Rashba term at the point $\boldsymbol{K}$, $(\mathbf{s} \times \mathbf{p}) \cdot \hat{z}$

$$h(\mathbf{K} + \mathbf{k}) = h(\mathbf{K}) = \sigma_x s_y - \sigma_y s_x$$

By TR invariance, we obtain the Hamiltonian at the point $\mathbf{K}' = -\mathbf{K}$:

$$h(\mathbf{K}') = s_y h^*(\mathbf{K}) s_y = -\sigma_x s_y^3 - \sigma_y s_y s_x s_y = -\sigma_x s_y - \sigma_y s_x$$

The Hamiltonians at $\mathbf{K}$ and $\mathbf{K}'$ can be written in compact form for both points $\mathbf{K}$ and $\mathbf{K}'$:

$$H_{Rashba} = \sigma_x \tau_z s_y - \sigma_y s_x$$

The Rashba term by itself does not open a full gap in the spectrum. This is due to the fact that this term does not anticommute with the Hamiltonian. At the point $\mathbf{K}$, we can add $\lambda_{Rashba}(s_y \sigma_x - s_x \sigma_y)$ to the $H_0(\mathbf{K} + \mathbf{k})$ Hamiltonian and get the four eigenvalues

$$-\lambda_{Rashba} \pm \sqrt{k^2 + \lambda_{Rashba}^2}, \qquad \lambda_{Rashba} \pm \sqrt{k^2 + \lambda_{Rashba}^2}$$

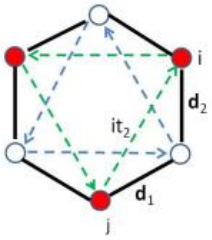
This shows that the Hamiltonian is gapless at the point $\mathbf{K}$ regardless of the value of the Rashba coupling. Out of four bands (two from the sublattice and two from spin), two are fully gapped and two are gapless. Even though it does not open a full gap by itself, the Rashba term can be used to influence (reduce or enhance) the Kane and Mele gap ($\sigma_z s_z \tau_z$) because that gap term anticommutes with the Rashba term.

Lattice Version

Consider the tight-binding Hamiltonian of graphene, which generalizes Haldane's model to include spin with TR-invariant spin-orbit interactions

$$H_{KM} = t \sum_{\langle ij \rangle} c_i^\dagger c_j + i\,\lambda_{\mathrm{SOC}} \sum_{\langle\langle ij \rangle\rangle} v_{ij} c_i^\dagger s_z c_j + i\lambda_{Rashba} \sum_{\langle ij \rangle} c_i^\dagger (\mathbf{s} \times \hat{\mathbf{d}}_{ij})_z c_j + \lambda_v \sum_i \xi_i\, c_i^\dagger c_j$$

The first term is the usual nearest neighbor (NN) hopping term on the honeycomb lattice, where we have suppressed the spin index on the electron operators. The second term is the mirror symmetric spin-orbit interaction which involves spin dependent next near neighbor (NNN) hopping. Here $c_i = (c_{i\uparrow}, c_{i\downarrow})$ and $v_{ij} = \frac{2}{\sqrt{3}}\left(\hat{\mathbf{d}}_1 \times \hat{\mathbf{d}}_2\right)_z = -v_{ji} = \pm 1$ are identical bond signs, depending on the orientation of the two nearest neighbor bonds. $\hat{\mathbf{d}}_1$ and $\hat{\mathbf{d}}_2$ are unit vectors along the two bonds the electron traverses in going from site j to i. $v_{ji} = +1/-1$ if the electron makes a left/right turn to get the second bond (see figure). s_z is a Pauli matrix describing the electron's spin. Specifically, the λ_{SOC} term does not violate mirror symmetry with respect to the lattice plane: if $z \to -z$, then the v_{ij} changes sign, but so does the spin s_z to keep the Hamiltonian unchanged. On the other hand, the Rashba term explicitly violates mirror symmetry. Let's look at the third term. The third term is a nearest neighbor Rashba term, which explicitly violates the $z \to -z$ mirror symmetry, and will arise due to a perpendicular electric field or interaction with a substrate. The fourth term is a staggered sublattice potential. The λ_v is the inversion-symmetry-breaking term -- $\xi_i = \pm 1$ depending on whether i is the *A* or *B* site, which we include to describe the transition between the QSH phase and the normal insulator. This term violates the symmetry under twofold rotations in the plane.



Likewise, for each **k** the Bloch wave function is a four-component eigenvector $|u(\mathbf{k})\rangle$ of the Bloch Hamiltonian matrix $\mathcal{H}(\mathbf{k})$. The 16 components of $\mathcal{H}(\mathbf{k})$ may be written in terms of the identity matrix, 5 Dirac matrices Γ^a and their 10 commutators $\Gamma^{ab} = [\Gamma^a, \Gamma^b]/(2i)$.

We choose the following representation of the Dirac matrices:

$$\Gamma^{(1,2,3,4,5)} = \left(\sigma_x I, \sigma_z I, \sigma_y s_x, \sigma_y s_y, \sigma_y s_z\right)$$

where the Pauli matrices σ_k and s_k represent the sublattice and spin indices. The Γ^{ab} can be easily obtained by commuting the Clifford generators but will not be written here. The Clifford matrices chosen by Kane and Mele are invariant under time reversal, such that the five Dirac matrices are even under $\mathcal{T}$: $\mathcal{T}\Gamma^a\mathcal{T}^{-1} = \Gamma^a$ (with $\mathcal{T} = i\left(I \otimes s_y\right)K$), while the 10 commutators are odd, $\mathcal{T}\Gamma^{ab}\mathcal{T}^{-1} = -\Gamma^{ab}$.

A generic 4 × 4 Hamiltonian (without any specific symmetries) has the form

$$\mathcal{H}(\mathbf{k}) = \sum_{a=1}^{5} d_a(\mathbf{k})\Gamma^a + \sum_{a<b=1}^{5} d_{ab}(\mathbf{k})\Gamma^{ab}$$

where the $d(\mathbf{k})$ are given below:

d_1	$t(1 + 2\cos x \cos y)$	d_{12}	$-2t \cos x \cos y$
d_2	λ_v	d_{15}	$\lambda_{\mathrm{SO}}(2\sin 2x - 4\sin x \cos y)$
d_3	$\lambda_{Rashba}(1 - \cos x \cos y)$	d_{23}	$-\lambda_{Rashba} \cos x \cos y$
d_4	$-\sqrt{3}\lambda_{Rashba} \sin x \sin y$	d_{24}	$\sqrt{3}\lambda_{Rashba} \sin x \cos y$

Note that $\mathcal{H}(\mathbf{k}+\mathbf{G})=\mathcal{H}(\mathbf{k})$ for reciprocal lattice vectors $\mathbf{G}$, so $\mathcal{H}(\mathbf{k})$ is defined on a torus. The $\mathcal{T}$ invariance of $\mathcal{H}$ is reflected in the symmetry (anti-symmetry) of d_a (d_{ab}) under $\mathbf{k}\to-\mathbf{k}$:

$$d_a(-\mathbf{k})=d_a(\mathbf{k}),\ \ d_{ab}(-\mathbf{k})=-d_{ab}(\mathbf{k})$$

The above $\mathcal{H}(\mathbf{k})$ gives four energy bands, two of which are occupied. For $\lambda_{Rashba}=0$, we saw this model is easy to diagonalize and gives two copies of the Haldane model. There is an energy gap with magnitude $\left|6\sqrt{3}\lambda_{\mathrm{SO}}-2\lambda_v\right|$. For $\lambda_v>3\sqrt{3}\lambda_{\mathrm{SO}}$ the system is an inversion-symmetry-breaking dominated phase and the gap is dominated by λ_v, and the system is an insulator. In contrast, $3\sqrt{3}\lambda_{\mathrm{SO}}>\lambda_v$ describes the QSH phase, in which the system has edge modes crossing the bulk gap, which were easily understood from the continuum argument previously given.

Though the Rashba term violates s_z conservation, for $\lambda_{Rashba}<2\sqrt{3}\lambda_{\mathrm{SO}}$ there is a finite region of the phase diagram that is adiabatically connected to the QSH phase at $\lambda_{Rashba}=0$. Both phases have a bulk energy gap and edge states, but in the QSH phase the edge states traverse the energy gap in pairs. At the transition between the two phases, the energy gap closes, allowing the edge states to ‘‘switch partners.’’ With $\lambda_{Rashba}=0$, this model has S_z conservation and we can easily talk about spin ↑ and spin ↓ sectors. With $\lambda_{Rashba}\neq0$, the S_z symmetry is broken, and we can no longer talk about two copies of fermions, one with spin ↑ and one with spin ↓, having separate Chern numbers. However, with $\lambda_{Rashba}\neq0$ but small so that we do not close the bulk gap, we observe that the edge modes do not disappear. In this case, edge modes traverse the bulk energy gap as long as λ_{Rashba} has not closed the bulk gap.

The behavior of the edge states signals a clear difference between the two phases. In the QSH phase for each energy in the bulk gap there is a single time reversed pair of eigenstates on each edge. Each edge has a pair of counterpropagating edge modes, which cross at some $\mathcal{T}$-invariant point. This crossing is protected by $\mathcal{T}$ symmetry. As long as time reversal is preserved, every *k*-point in the system must have a $\mathcal{T}$ orthogonal counterpart at $-\mathbf{k}$. The $\mathcal{T}$-invariant points must each have two states, by Kramers’ theorem. That means the gap can never open for a single pair of counterpropagating modes. For every edge in the system, we will observe a pair of edge modes. The dispersions of different edges can cross at some non-$\mathcal{T}$-invariant point, as in fig. 9.1(a), but the matrix elements coupling these points are zero because they have to fully traverse the system length. Hence there is no one-body, $\mathcal{T}$-invariant local perturbation term that can couple the two branches of a single pair of edge modes. Since $\mathcal{T}$ symmetry prevents the mixing of Kramers’ doublets these edge states are robust against small perturbations. The gapless states thus persist even if the spatial symmetry is further reduced [for instance, by removing the C_3 rotational symmetry in the tight-binding model]. Moreover, weak disorder will not lead to localization of the edge states because single particle elastic backscattering is forbidden.

This situation is different for an even number of a pair of edge modes on the same edge. In the insulating state the edge states do not traverse the gap. It is possible that for certain edge potentials the edge states in right Figure could dip below the band edge, reducing—or even eliminating—the edge gap. However, this is still distinct from the QSH phase because there will necessarily be an even number of Kramers’ pairs at each energy. In this case, the modes can open gaps at the intersection points between the two pairs (which can be away from a TRIM point). $\mathcal{T}$ does not

protect that gap opening because a gap opening at $k \neq 0$ can be accompanied by another gap opening at $-k \neq 0$, which can cancel the $\mathcal{T}$ breaking of the first. Different pairs of the two modes therefore scatter between themselves and can open a gap because the matrix elements are on the same edge. This allows elastic backscattering, so that these edge states will in general be localized by weak disorder. The QSH phase is thus distinguished from the simple insulator by the number of edge state pairs modulo 2.

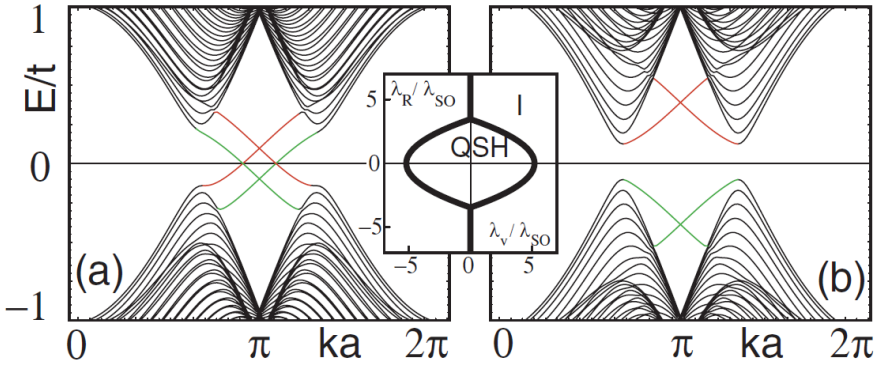


FIG. 1 (color online). Energy bands for a one-dimensional "zigzag" strip in the (a) QSH phase $\lambda_v = 0.1t$ and (b) the insulating phase $\lambda_v = 0.4t$. In both cases $\lambda_{SO} = .06t$ and $\lambda_R = .05t$. The edge states on a given edge cross at $ka = \pi$. The inset shows the phase diagram as a function of λ_v and λ_R for $0 <$

The above model was the first TRI topological insulator proposed by Kane and Mele in 2005. As it is TRI, the total Chern number cannot be non-zero. However, in the above Hamiltonian, the spin projections $|\uparrow\rangle, |\downarrow\rangle$ are good eigenstates. In the limit where the spin-↑ and spin-↓ blocks remain decoupled, we can define two winding numbers, $\mathcal{C}_\uparrow, \mathcal{C}_\downarrow$. Because time reversal is preserved, we have $\mathcal{C}_\uparrow = -\mathcal{C}_\downarrow$ and therefore $\mathcal{C}_\uparrow + \mathcal{C}_\downarrow = 0$. However, in this decoupled limit, we can use the spin-Chern number $\mathcal{C}_\sigma$ in each spin-sector to characterize the phases. Indeed

$$\nu = \frac{\mathcal{C}_\uparrow - \mathcal{C}_\downarrow}{2} \bmod 2 \ \in \mathbb{Z}_2$$

defines a good topological index $\mathbb{Z}_2$. The addition of a Rashba term $H_{Rashba} = \sigma_x \tau_z s_y - \sigma_y s_x$ removes this conserved quantity. While H_{Rashba} does not open a gap by itself, it can influence the SOC induced gap. However, the above topological index ν is not well defined anymore. In the following section we aim at deriving a $\mathbb{Z}_2$ index which does not rely on spin-Chern numbers.

$\mathbb{Z}_2$ index

Bulk Boundary Correspondence

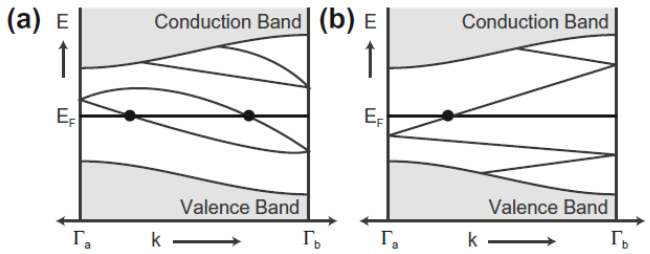


FIGURE 11 Electronic dispersion between two boundary Kramers degenerate points. In (a) the number of surface states crossing the Fermi energy E_F is even, whereas in (b) it is odd. An odd number of crossings leads to topologically protected metallic boundary states.

In Fig. 11 we show the electronic states associated with the edge of a $\mathcal{T}$ invariant 2D insulator as a function of the crystal momentum along the edge. Only half of the Brillouin zone $\Gamma_a = 0 < k_x < \Gamma_b = \pi/a$ is shown because $\mathcal{T}$ symmetry requires that the other half $-\pi/a < k < 0$ is a mirror image. The shaded regions depict the bulk conduction and valence bands separated by an energy gap. Depending on the details of the Hamiltonian near the edge there may or may not be states bound to the edge inside the gap. If they are present, however, then Kramers' theorem requires they be twofold degenerate at the $\mathcal{T}$ invariant momenta $k_x = 0$ and $k_x = \pi/a$ (which is the same as $-\pi/a$). Away from these special points, labeled $\Gamma_{a,b}$ in Fig. 11, a spin orbit interaction will split the degeneracy. There are two ways the states can connect. In Fig. 11(a) they connect pairwise. In this case the edge states can be eliminated by pushing all the bound states out of the gap. Between $k_x = 0$ and $k_x = \pi/a$, the bands intersect E_F an even number of times. In contrast, in Fig. 11(b) the edge states cannot be eliminated. The bands intersect E_F an odd number of times.

Which of these alternatives occurs depends on the topological class of the bulk band structure. Since each band intersecting E_F at k_x has a Kramers partner at $-k_x$, the bulk-boundary correspondence relates the number N_K of Kramers pairs of edge modes intersecting E_F to the change in the $\mathbb{Z}_2$ invariants across the interface,

$$N_K = \Delta\nu \mod 2$$

Why $\mathbb{Z}_2$?

Let's check what will happen to the edge states, when we include extra terms into the Hamiltonian.

Case I: 1 pair of edge states.

Let's consider the Kane-Mele model (or any other model with one pair of edge states). Here our spin up electron has C=1, so it has one left moving edge states $\epsilon = v_F k$. Spin down electrons have C=-1, so the edge states for spin down electrons are right moving $\epsilon = -v_F k$. We can write them together in one Hamiltonian as a 2×2 matrix.

$$\mathcal{H}(k) = \begin{pmatrix} v_F k & 0 \\ 0 & -v_F k \end{pmatrix} = v_F k \sigma_z$$

Introducing extra terms in the bulk resulting in extra terms in the edge states Hamiltonian.

Q: Can one add any extra term to open a gap for the edge states?

This question is important. If we can open the gap for the edge states, we can move the chemical potential into the gap and the edge states turns into an insulator. If the system has an insulating edge, it is a trivial insulator (edge states are no long topologically protect). But if we can prove that no matter what we do, we can NOT open a gap, the edge is always metallic, so we have a topologically protected edge state (topological insulator).

A: No, if we preserve the time-reversal symmetry. Yes, if we break the time-reversal symmetry. In other words, if the system has time-reversal symmetry, it is a topological state with protected edge states. But if we break the time-reversal symmetry, it is a trivial insulator.

First, let's break the time-reversal symmetry, by adding a $m\sigma_z$ term to the Hamiltonian and see what will happen. This term breaks the time-reversal symmetry. $\mathcal{T}\sigma \to -\sigma$. So $\mathcal{T}m\sigma_z \to -m\sigma_z$

$$\mathcal{H}(k) = \begin{pmatrix} v_F k & m \\ m & -v_F k \end{pmatrix}$$

Dispersion:

$$\epsilon_{\pm} = \sqrt{(v_F k)^2 + m^2}$$

We have a gap in the spectrum of the edge state. The edge state is NOT gapless. If we have chemical potential inside the gap, the edge is NOT a metal. (the system is not a topological insulator).

Second, let's preserve the time-reversal symmetry and see what will happen. If we preserve the time-reversal symmetry, the most general Hamiltonian we can write down is:

$$\mathcal{H}(k) = \sum\nolimits_{c_{n,i}} k^{2n+1}\sigma_i + \sum b_n k^{2n} I$$

At $k = 0$,

$$\mathcal{H}(k = 0) = b_0 I = \begin{pmatrix} b_0 & 0 \\ 0 & b_0 \end{pmatrix}$$

Eigenvalues at $k = 0$: $\epsilon_{\pm} = b_0$. Gap at $k = 0$: $\Delta(k = 0) = \epsilon_+ - \epsilon_- = 0$

No gap! No matter what chemical potential we choose, the system must be a metal at the edge. (a topological insulator).

Case II: 2 pairs of edge states.

Assume the spin up state is in a quantum Hall state with C=2, and the spin down state has C=-2. We will have two left moving electrons with spin up, and two right moving electrons with spin down.

$$H = \sum_k \begin{pmatrix} c_{k\uparrow}^\dagger & d_{k\uparrow}^\dagger & c_{k\downarrow}^\dagger & d_{k\downarrow}^\dagger \end{pmatrix} \mathcal{H} \begin{pmatrix} c_{k\uparrow}^\dagger \\ d_{k\uparrow}^\dagger \\ c_{k\downarrow}^\dagger \\ d_{k\downarrow}^\dagger \end{pmatrix}$$

$$\mathcal{H}(k) = \begin{pmatrix} v_{F1}k & 0 & 0 & 0 \\ 0 & v_{F2}k & 0 & 0 \\ 0 & 0 & -v_{F1}k & 0 \\ 0 & 0 & 0 & -v_{F2}k \end{pmatrix}$$

Now, one can add one extra term,

$$\mathcal{H}(k) = \begin{pmatrix} v_{F1}k & 0 & 0 & m \\ 0 & v_{F2}k & -m & 0 \\ 0 & -m & -v_{F1}k & 0 \\ m & 0 & 0 & -v_{F2}k \end{pmatrix}$$

This Hamiltonian is time-reversal invariant. As we will show in the next lecture, the time-reversal operator is

$$U_T = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

Under time-reversal,

$$\mathcal{H}(k) \to \mathcal{H}'(k) = (U_T \mathcal{H}(k) U_T^{-1})^*$$

Here * means complex conjugate. It can be easily checked that the $\mathcal{H}(k)$ we wrote above is invariant under time-reversal

$$\mathcal{H}(k) \to \mathcal{H}'(k) = (U_T \mathcal{H}(k) U_T^{-1})^* = \mathcal{H}(k)$$

Here, we consider a simple case with $v_{F1} = v_{F2}$ (for $v_{F1} \neq v_{F2}$, all qualitative results remain the same). The eigenvalues are:

Two degenerate upper band: $\epsilon = \sqrt{m^2 + k^2 v_{F1}^2}$ and two degenerate lower band $\epsilon = -\sqrt{m^2 + k^2 v_{F1}^2}$. There is a gap $\Delta = 2m$.

Generic cases

In a time-reversal invariant 2D system:

- If we have (2*n*+1) pairs of edge modes, one can add extra terms to gap 2*n* of them. But there will always be 1 pair which is gapless (topological insulator)

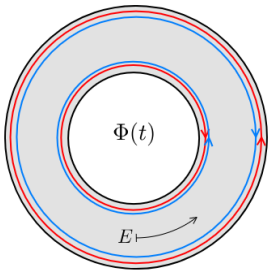
- If we have 2*n* pairs of edge modes, one can always add some extra term to gap all of them. (trivial insulator)

These two classes correspond to the simplest non-trivial set $\mathbb{Z}_2$ (the set of integers modulo 2). Thus 2D topological insulators are characterized by a $\mathbb{Z}_2$ quantum number ν that takes two possible values, namely 0 and 1, or, equivalently, even and odd, corresponding to the trivial and non-trivial states, respectively.

Connection to high energy, one pair of edge states give us a Weyl fermion. *n* pairs of edge states give us *n* Weyl fermions. If we enforce chiral symmetry, A Weyl fermion can only get mass (open gap) but two Weyl fermions can get mass, by combining them into a Dirac fermion. If we have 2*n* Weyl fermions, we can pair them up into *n* Dirac fermions, and give all of them some mass. But if we have 2*n*+1 Weyl fermions, we cannot do so. We can pair 2*n* of them into *n* Dirac fermions and give them mass, but there will be one left, which is massless (i.e. gapless).

Charge polarization in 1D

Since the time-reversal invariant topological insulator is two dimensional like a quantum Hall system, it is reasonable to put the system in a Corbino geometry and change the flux through the system, creating an azimuthal electric field as shown in the right figure.



However, because of time-reversal symmetry, the system is forbidden from having a Hall conductance and therefore there cannot be any charge transfer between the two edges of the disk. For instance, if we consider two copies of the Haldane model with opposite spin, there will be two quantum Hall pumps working in opposite directions (one transferring charge from the inner edge to the outer edge, the other one from the outer edge to the inner one). So the net charge transferred is zero.

Because the two pumps act on electrons with opposite spin, you might be tempted to define a spin current, which would flow in response to the electric field, orthogonal to it. However, as we just discussed, the spin along a given direction may not be conserved, so generally this is not a good way to define a robust pumping effect.

We consider a one dimensional system on a lattice (with lattice constant a = 1) with periodic Bloch wave functions

$$|\psi_{nk}\rangle = \frac{1}{\sqrt{L}} e^{ikx} |\varphi_{nk}(x)\rangle \quad \text{with} \quad |\varphi_{nk}(x)\rangle = |\varphi_{nk}(x+1)\rangle$$

The corresponding Wannier functions are defined as

$$\left|W_{R,n}(x)\right\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} dk\, e^{-ik(R-x)} \left|\varphi_{n,k}(x)\right\rangle$$

with R is a lattice vector. Note that the Wannier functions are not gauge invariant as the relative phase between different $\left|\varphi_{n,k}(x)\right\rangle$ is not a priori fixed. For a smooth gauge, the Wannier functions are exponentially localized around a well-defined center.

The total charge polarization for this 1D system is now given by the sum over all the bands of the center of charge of the Wannier states

$$P = \sum_{n\ filled} \int dx \left\langle W_{0,n}(x)\middle|x\middle|W_{0,n}(x)\right\rangle$$

We try to write this polarization in a more familiar way

$$\begin{aligned} P &= \sum_{n\ filled} \frac{1}{(2\pi)^2}\frac{1}{L}\int_{-\pi}^{\pi} dk_1 \int_{-\pi}^{\pi} dk_2 e^{i(k_2-k_1)x} \left\langle \varphi_{n,k_1}\middle|i\partial_k\middle|\varphi_{n,k_2}\right\rangle \\ &= \sum_{n\ filled} \frac{i}{2\pi}\int_{-\pi}^{\pi} dk \left\langle \varphi_{nk}|i\partial_k|\varphi_{nk}\right\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} dk\, \mathcal{A}_x(k) \end{aligned}$$

with

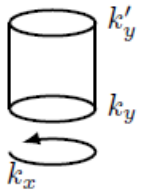
$$\mathcal{A}_x(k) = \sum_{n\ filled} i\left\langle \varphi_{n,k}\middle|i\partial_k\middle|\varphi_{n,k}\right\rangle$$

Two comments are in order:

1. If we re-gauge $|\varphi\rangle \to e^{i\vartheta(k)}|\varphi\rangle$ with a $\vartheta(k)$ that is winding by $2\pi m$ throughout the Brillouin zone, the corresponding polarization changes to
   $$P \to P + m$$
   This is ok, as charge polarization is anyway only defined up to a lattice vector in a translational invariant system, as shown by $P = \sum_{n\ filled} \int dx \left\langle W_{0,n}(x)\middle|x\middle|W_{0,n}(x)\right\rangle$.
2. Although P depends on the chosen gauge, but changes in P by an adiabatic and continuous change in system parameters are gauge independent. So let us imagine a tuning parameter k_y with
   $$H(k_y) \to H(k_y')$$

which is slow in time. The change in charge polarization is given by

$$\Delta P = P(k_y') - P(k_y) = \frac{1}{2\pi}\left[\int_{-\pi}^{\pi} dk\, \mathcal{A}(k, k_y) - \int_{-\pi}^{\pi} dk\, \mathcal{A}(k, k_y')\right]$$



If we use Stokes' theorem we arrive at

$$\Delta P = \int_{k_y}^{k_y'} dk_y \int_{-\pi}^{\pi} dk\, \mathcal{F}(k, k_y)$$

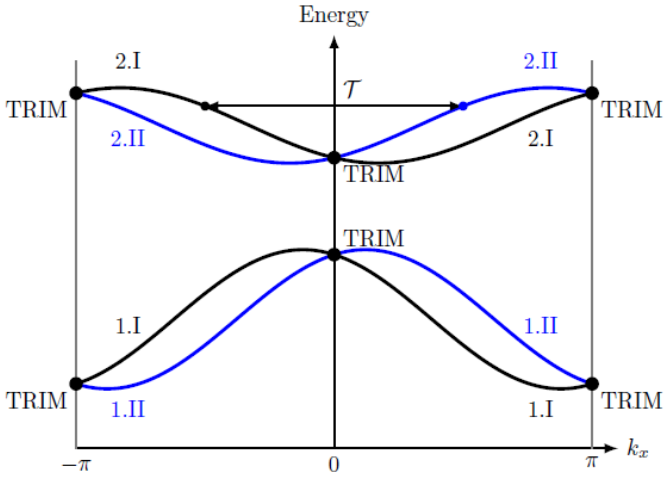
where $\mathcal{F}(k, k_y) = i\sum_n \left(\left\langle \partial_{k_y}\varphi_{nk} \middle| \partial_{k_x}\varphi_{nk} \right\rangle\right)$. By choosing $k_y' = k_y + 2\pi$ to close the cylinder into a torus, we find for the change in charge polarization is given by the integral over the torus, $\Delta P = C$ where C is the Chern number. It characterizes the charge difference that is pumped in each cycle. We know, however that $\mathcal{F}(-k, -k_y) = -\mathcal{F}(k, k_y)$ for a TRI system and hence the Chern number is equal to zero for TRI systems.

We have so far defined the charge polarization, which is the sum of the Berry curvature integrated over the BZ. The difference in the charge polarization over the full BZ is the Hall conductance.

Time Reversal Polarization

Let us now try to generalize the charge pumping approach to the TRI setup. For this it is beneficial to look at the structure of a generic energy diagram as shown in Fig below. Under time-reversal, momenta k are mapped to $-k$.

Owing to the symmetry between k and $-k$ we can constrain ourselves to only half the Brillouin zone. In this half, we label all bands by 1.I, 1.II, 2.I, 2.II, ... The Arabic number simply label pairs of bands. Due to the double degeneracy at the TRIMs, we need an additional (Roman number) to label the two (sub)-bands emerging form the TRIMs. Time reversal transforms eigenstates at k of bands I into eigenstates at $-k$ of bands II, and vice versa, but only up to a phase factor



$$|\varphi_{n.\mathrm{I}}(k)\rangle = e^{i\chi_{n,k}}\mathcal{T}|\varphi_{n.\mathrm{II}}(k)\rangle$$

The TR operator acting on this then gives

$$\mathcal{T}|\varphi_{n.\mathrm{I}}(k)\rangle = \mathcal{T}e^{i\chi_{n,k}}\mathcal{T}|\varphi_{n.\mathrm{II}}(k)\rangle = e^{-i\chi_{n,k}}\mathcal{T}^2|\varphi_{n.\mathrm{II}}(k)\rangle = -e^{-i\chi_{n,k}}|\varphi_{n.\mathrm{II}}(k)\rangle$$

Therefore

$$|\varphi_{n.\mathrm{II}}(k)\rangle = -e^{i\chi_{n,k}}\mathcal{T}|\varphi_{n.\mathrm{I}}(k)\rangle$$

We now try to construct the polarization for only one of the two labels (TR pair) $\alpha =$ I or II

$$P^{\alpha} = \frac{1}{2\pi}\int_{-\pi}^{\pi} dk\, \mathcal{A}^{\alpha}(k)$$

where the Berry connection is locally defined as

$$\mathcal{A}^{\alpha}(k) = i \sum_{n\ filled} \langle\varphi_{n.\alpha}(k)|\partial_k|\varphi_{n.\alpha}(k)\rangle$$

Using the combinations of the partial polarizations, one further defines the charge polarization P^{ρ} and the time reversal polarization $P^{\mathcal{T}}$. It is clear that

$$P^{\rho} = P^{\mathrm{I}} + P^{\mathrm{II}}$$

will always vanish since for time reversal invariant topological insulators. However, the same must not hold for the time reversal polarization

$$P^{\mathcal{T}} = P^{\mathrm{I}} - P^{\mathrm{II}}$$

The time reversal polarization delivers the topological $\mathbb{Z}_2$ invariant. The problem is, that we assigned the labels I and II. It is not a priori clear if this can be done in a gauge invariant fashion. There is a way however, to formulate the same $\mathcal{T}$-polarization $P^{\mathcal{T}}$ in a way that does not rely on a specific labeling of the Kramers pairs. This can be achieved by the use of the so-called sewing matrix

$$B_{mn}(k) = \langle\varphi_m(-k)|\mathcal{T}|\varphi_n(k)\rangle$$

$B(k)$ has the following properties: (i) it is unitary, and (ii) it is anti-symmetric, i.e., $B^T(k) = -B(k)$ only if k is a TRIM, as shown in the following:

We can label

$$B^{\mathrm{II,I}}_{mn}(k) = \langle\varphi^{\mathrm{II}}_m(-k)|\mathcal{T}|\varphi^{\mathrm{I}}_n(k)\rangle = -\delta_{mn}e^{-i\chi_{n,-k}}$$

$$B^{\mathrm{I,II}}_{mn}(k) = -B^{\mathrm{II,I}}_{mn}(-k)$$

The matrix $B^{\mathrm{I,II}}_{mn}$ is the sewing matrix between the Kramers' pair bands I and II. There could be multiple such Kramers' pairs, indexed by m = 1, 2, . . . , N, but the sewing matrix couples the

Kramers' pairs only two by two. Hence, the matrix B becomes a block-diagonal matrix of the subblock matrices:

$$\begin{bmatrix} 0 & -e^{-i\chi_{n,-k}} \\ e^{-i\chi_{n,k}} & 0 \end{bmatrix}$$

We see that irrespective of $\chi_{n,k}$ at $k = 0, \pi$ only, the preceding matrix is antisymmetric.

Using this matrix and tedious computation, one can show that the TR polarization is

$$P^{\mathcal{T}} = \frac{1}{i\pi} \log \left[\frac{\sqrt{\det B(\pi)}}{\mathrm{Pf}\, B(\pi)} \frac{\mathrm{Pf}\, B(0)}{\sqrt{\det B(0)}} \right]$$

This expression is manifestly invariant under SU(2n) rotations within the filled bands. Moreover, it only depends on the two TRIMs $k = 0, \pi$, and it is defined modulo two. This TR polarization is a nontrivial $\mathbb{Z}_2$ index characterizing a TR insulator.

The TR polarization is defined for a 1-D system: it tells how the difference in the polarization of the elements of a Kramers' pair behave in a 1-D system. It depends only on the sewing matrices at the points 0 and 2π. In the case of the Hall conductance, in order to define the Chern number as an invariant in a pumping cycle, we looked at how the charge polarization varied in a pumping cycle from 0 to a period T when a periodic parameter was varied. In anticipation of the fact that the periodic parameter is another momentum, we can also denote this periodic parameter by k_y, the momentum in another direction. In a pumping cycle (between 0 and T—or between 0 and 2π if we are to talk about k_y), a TR Hamiltonian occurs at $t = 0$ and $t = T/2$ (or $k_y = 0, \pi$). This is clear because only these two points map into themselves under time reversal. When we consider the evolution of the Hamiltonian through the cycle, the change in TR polarization in half a cycle defines a $\mathbb{Z}_2$ topological invariant.

The $\mathbb{Z}_2$ invariant is related to the center of the occupied Wannier orbitals as a function of t.

Let us now see how we can describe changes in the time-reversal polarization under the influence of an additional parameter k_y. Written as in the expression of $P^{\mathcal{T}}$, it is only defined for $k_y = 0, \pi, 2\pi$, i.e, at TRIMs. In Fig. 5.3 we illustrate what we can expect from such a smooth change. We start at $k_y = 0$. If we now change k_y slowly, we know that due to TRI, we cannot build up a charge polarization. However, the Wannier centers of two Kramers pairs will evolve in opposite direction. At $k_y = \pi$, we can check how far these centers evolved away from each other. As $P^{\mathcal{T}}$ is well defined and equal to 0 or 1 we have two options: (i) Each Wannier center meets up with one coming from a neighboring site [Fig. 5.3(a)]. This gives rise to $P^{\mathcal{T}}(k_y = \pi) = 1$ and this

effect is called pair switching. (ii) The centers fall back onto each other again [Fig. 5.3(b)], resulting in $P^{\mathcal{T}}(k_y = \pi) = 0$.

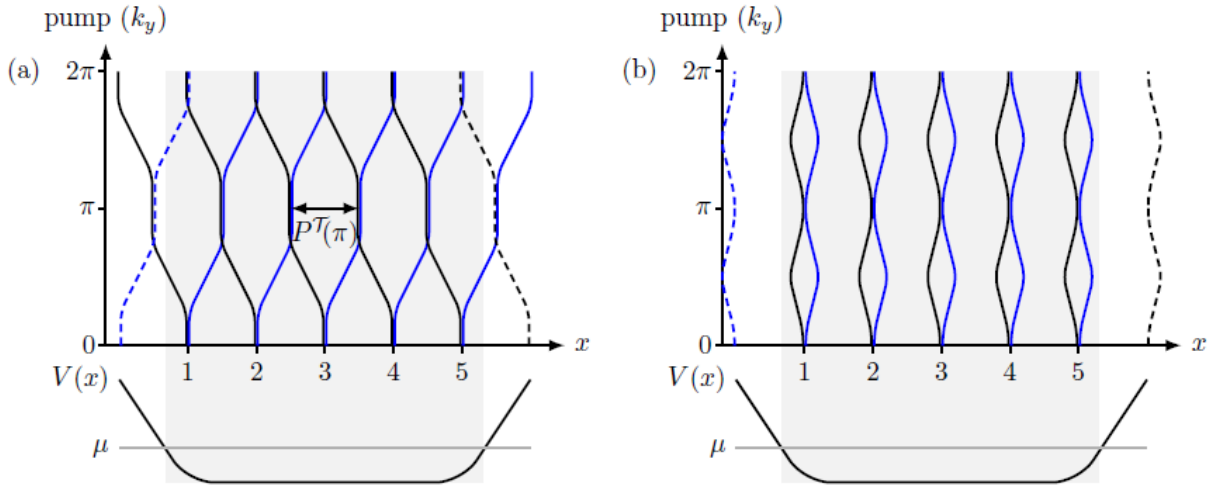


Figure 5.3: (a) Pumping of time reversal polarization in a topologically non-trivial state. (b) Pumping of time reversal polarization in a trivial state.

Let us further assume that we have a smooth confining potential $V(x)$ in x-direction. As in the case of the quantum Hall effect, we see how states can be pushed up-hill or pulled down-hill as a function of k_y. However, as opposed to the quantum Hall effect, we have here the situation that on each edge we have both a state coming down in energy as well as one climbing up! From that we conclude that if we have pair-switching, we expect two counter-propagating edge states on both sides of the sample.

(1) At t = 0, T/2, TR symmetry requires that the Wannier states come in TR pairs—the center of the occupied Wannier orbitals for each partner (the eigenvalues of the position operator) must be equal.

(2) However, in going from t = 0 to t = T/2, if the system is in a $\mathbb{Z}_2$ nontrivial topological state, the Wannier states "switch partners," as in the figure. While doing this, the TR polarization changes by 1 as it records the difference between the positions of the TR Wannier states.

(3) The partner switching has a nice explanation in terms of a system with edges: put the system on a cylinder with length in the y-direction of just a single unit cell. If we start with a state of all Wannier-occupied orbitals at t = 0, once we insert a flux of π (to go to T/2 or $k_y = \pi$), we see that one of the occupied Wannier orbitals closest to the edge has switched partners: its new, degenerate partner at T/2 is a previously unoccupied orbital.

(4) In the bulk, the partner switching does not result in much, except for the relabeling of the Wannier orbitals. On the edge, however, partner switching results in the occupied Wannier orbital at $t = 0$ switching partners and having as a partner at $t = T/2$, a formerly (before applying the flux π) unoccupied Wannier orbital. This means that the partner switching results in the appearance of an unoccupied Wannier orbital at each end (edge) of the sample.

(5) As such, the ground-state degeneracy will have changed: on one edge, the TR partner switching at $t = T/2$ results in an unoccupied Wannier orbital coming down and becoming degenerate with one of the $t = 0$ occupied Wannier orbitals.

(6) Consequently, the state at $t = T/2$ is twofold degenerate on one edge, being a doublet (we have two states and occupation number 1—the occupied Wannier orbital at $t = 0$—which means we can occupy either of the two states at $t = T/2$). The total degeneracy, when we have two edges, is four, two from one edge and two from the other.

(7) We can see this mode clearly by again solving the Hamiltonian for a cylinder, where the length of the cylinder in the translationally invariant direction (along the edge) equals one lattice constant. In this case, the spectrum of states as a function of the momentum parallel to the surface is identical to the spectrum of states as a function of the flux added to the system.

There is in one-to-one connection between the TR polarization and the appearance of edge modes. If the periodic parameter t represents the k_y momentum, it means that, once a surface is placed perpendicular to the x-direction, we can project the two bulk TR-invariant momenta $k_x = 0, \pi$ at $t = 0$ into one point denoted by $k_y = 0$ on the surface, while the other two bulk TR-invariant momenta $k_x = 0, \pi$ at $t = T/2$ go into one point on the surface $k_y = \pi$. If the difference between the TR polarization at $t = 0$ and the TR polarization at $t = T/2$ is unity, then there is a mismatch, which cannot be realized in a gapped system connected by an edge to the vacuum: an odd number of edge modes crossing the Fermi level between $t = 0$ and $t = T/2$ has to exist. Hence, the topological invariant corresponds to an edge (or edge states) of the effective Brillouin zone, which is a fundamental domain for the time reversal $\mathbb{Z}_2$ action. One can compute

$$P^{\mathcal{T}} = \frac{1}{2\pi}\int_{-\pi}^{\pi} dk\,(\mathcal{A}^{I} - \mathcal{A}^{II}) = \frac{1}{2\pi i}\left[\ln\frac{\det\big(B(\pi)\big)}{\det\big(B(0)\big)} - 2\ln\frac{\mathrm{Pf}\,\big(B(\pi)\big)}{\mathrm{Pf}\,\big(B(0)\big)}\right]$$

After exponentiating both sides and using the identity $e^{i\pi} = -1$, one obtains

$$(-1)^{P^{\mathcal{T}}} = \prod_{\Lambda_l = 0,\pi} \frac{\sqrt{\det B(\Lambda_l)}}{\mathrm{Pf}\,B(\Lambda_l)} = \prod_{\Lambda_l = 0,\pi} \mathrm{sign}\big(\mathrm{Pf}\,B(\Lambda_l)\big)$$

where Λ_l are the TR-invariant momenta. This is the $\mathbb{Z}_2$ topological invariant in two dimensions. It is important to point out that $\sqrt{\det B(\Lambda_l)}$ is a continuous choice of square root defined on the interval from 0 to π. This is possible by extending a choice of a branch of the square root at 0 along the interval $[0, \pi]$. This is what gives the coherent choice of normalization mentioned above, which

manifests itself as a choice of the sign of the Pfaffian. Since both the Pfaffian and the chosen branch are square roots of the determinant, they differ only by a sign ±1.

Compute the TR polarization for each of those edge points—it will be the product of the $\mathrm{Pf}\, B/\sqrt{\det B}$ for the two bulk points that project onto the edge TR point. If the TR polarization changes sign in between the points, then there must be an edge mode crossing any Fermi level (situated in the gap) between the edge TR-invariant points. The edge mode crossing the bulk gap will give the TR-polarization change of 1.